UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P O Box 1450
Alexandria, Virginia 22313-1450
www uspto gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/649,382	08/27/2003	Nebojsa Jojic	301911.01	6064

27662 7590 05/20/2008

MICROSOFT CORPORATION
C/O LYON & HARR, LLP
300 ESPLANADE DRIVE
SUITE 800
OXNARD, CA 93036

EXAMINER	
RASHID, DAVID	
ART UNIT	PAPER NUMBER
2624	
MAIL DATE	DELIVERY MODE
05/20/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/649,382	JOJIC ET AL.	
	Examiner	Art Unit	
	DAVID P. RASHID	2624	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *2/25/2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-32* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-32* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

[1] All of the examiner's suggestions presented herein below have been assumed for examination purposes, unless otherwise noted.

Amendments

[2] This office action is responsive to the claim and specification amendment received on February 25, 2008. Claims 1-32 remain pending.

Specification

[3] In response to applicant's specification amendments and remarks received on February 25, 2008, the previous specification objections are withdrawn.

Claim Rejections - 35 USC § 101

[4] 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

[5] The USPTO "Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility" (Official Gazette notice of 22 November 2005), Annex IV, reads as follows:

> Descriptive material can be characterized as either "functional descriptive material" or "nonfunctional descriptive material." In this context, "functional descriptive material" consists of data structures and computer programs which impart functionality when employed as a computer component. (The definition of "data structure" is "a physical or logical relationship among data items, designed to support specific data manipulation functions." The New IEEE Standard Dictionary of Electrical and Electronics Terms 308 (5th ed. 1993).) "Nonfunctional descriptive material" includes but is not limited to music, literary works and a compilation or mere arrangement of data.
>
> When functional descriptive material is recorded on some computer-readable medium it becomes structurally and functionally interrelated to the medium and will be statutory in most cases since use of technology permits the function of the descriptive material to be realized. Compare In re Lowry, 32 F.3d 1579, 1583-84, 32 USPQ2d 1031, 1035 (Fed. Cir. 1994) (claim to data structure stored on a computer readable medium that increases computer efficiency held statutory) and Warmerdam, 33 F.3d at 1360-61, 31 USPQ2d at 1759 (claim to computer having a specific data structure stored in memory held statutory product-by-process claim) with Warmerdam, 33 F.3d at 1361, 31 USPQ2d at 1760 (claim to a data structure per se held nonstatutory).
>
> In contrast, a claimed computer-readable medium encoded with a computer program is a computer item which defines structural and functional interrelationships between the computer program and the rest of the computer

> which permit the computer program's functionality to be realized, and is thus statutory. See Lowry, 32 F.3d at 1583-84, 32 USPQ2d at 1035.

[6] **Claims 1-22** and **23-32** are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter as follows. **Claims 1-22** define "[a] system…comprising using a computer to…" and **claims 23-32** define "[a] computer-implemented process…comprising using a computer to…" embodying functional descriptive material. However, the claim does not define a computer-readable medium or memory and is thus non-statutory for that reason (i.e., "When functional descriptive material is recorded on some computer-readable medium it becomes structurally and functionally interrelated to the medium and will be statutory in most cases since use of technology permits the function of the descriptive material to be realized"-Guidelines Annex IV). That is, the scope of the presently claimed "system… comprising using a computer to…" and "computer-implemented process…comprising using a computer to…" can range from paper on which the program is written, to a program simply contemplated and memorized by a person. The examiner suggests amending the claim to embody the program on "A system for automatically decomposing an image sequence, comprising a computer-storage media storing a program such that when executed perform the following process actions…" or equivalent in order to make the claim statutory. Any amendment to the claim should be commensurate with its corresponding disclosure.

[7] The applicant must also note normally a claim would be statutory when residing on a "computer-readable medium" or its definite equivalent (e.g. "computer-readable media"). However, the specification, at pages 11 - 12 defines the claimed computer readable medium as encompassing statutory media such as a "ROM", "hard drive", "optical drive", etc, as well as ***non-***

statutory subject mater such as a "carrier wave", "modulated data signal", and other equivalents thereof.

A "signal" embodying functional descriptive material is neither a process nor a product (i.e., a tangible "thing") and therefore does not fall within one of the four statutory classes of § 101. Rather, "signal" is a form of energy, in the absence of any physical structure or tangible material.

Because the full scope of the claim as properly read in light of the disclosure encompasses non-statutory subject matter, the claim as a whole is non-statutory. The examiner suggests amending the claim to *include* the disclosed tangible computer readable media, while at the same time *excluding* the intangible media such as signals, carrier waves, etc. Any amendment to the claim should be commensurate with its corresponding disclosure.

Claim Rejections - 35 USC § 102

[8] The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

[9] **Claims 1-3, 5-6, 14, 18-19,** and **23-24** are rejected under 35 U.S.C. 102(b) as being anticipated by Foote et al. (US 6,404,925 B1).

Regarding **claim 1**, Foote discloses a system (fig. 1; fig. 2) for automatically decomposing an image sequence (fig. 2, item 201), comprising using a computer (computer in fig. 1) to perform the following process actions:

providing an image sequence (fig. 2, item 201) of at least one image frame (fig. 3, items 301-308) of a scene;

providing a preferred number of classes of objects ("pre-defined set of classes" in 5:14-16 wherein the "pre-defined set of classes" were a preferred number provided at some point in time) to be identified within the image sequence;

automatically decomposing the image sequence into the preferred number of classes of objects in near real-time ("segmenting...into a pre-defined set of classes" in 5:14-16 is an act of "decomposing")

Regarding **claim 2**, Foote discloses the system of claim 1 wherein providing the preferred number of objects ("pre-defined set of classes" in 5:14-16) comprises specifying the preferred number of classes of objects via a user interface (a user interface is visual interface from which a user can interact with such as fig. 22; a pre-defined set of classes suggests that some sort of user interface must have been used to "define" the set of classes; "[t]he feature used for classification are general, so that users can define arbitrary class types" in 5:18-20).

Regarding **claim 3**, Foote discloses the system of claim 1 wherein decomposing the image sequence (fig. 2, item 201) into the preferred number of objects ("segmenting...into a pre-defined set of classes" in 5:14-16) comprises automatically learning a 2-dimensional model (fig. 3, items 310-322) of each object class (7:13-15).

Regarding **claim 5**, Foote discloses the system of claim 1 wherein automatically decomposing the image sequence (fig. 2, item 201) into the preferred number of object classes ("pre-defined set of classes" in 5:14-16) comprises performing an inferential probabilistic analysis (fig. 2, items 202-205; "Gaussian distributions" in 5, line 65-6, line 2) of each image frame for identifying ("segmenting…into a pre-defined set of classes" in 5:14-16) the preferred number of object class appearances within the image sequence.

Regarding **claim 6**, Foote discloses the system of claim 5 wherein performing an inferential probabilistic analysis of each image frame comprises performing a variational generalized expectation-maximization analysis (21:55-62) of each image frame (fig. 3, items 301-308) of the image sequence (fig. 2, item 201), wherein the expectation-maximization analysis employs a Viterbi algorithm (6:43-45; 16:40-42) in a process of filling in values of hidden variables (21:55-62; variables in fig. 4) in a model describing the object class.

Regarding **claim 14**, Foote discloses the system of claim 1 wherein automatically decomposing the image sequence into the preferred number of object classes comprises performing a probabilistic variational expectation-maximization analysis (21:55-62).

Regarding **claim 18**, Foote discloses the system of claim 1 further comprising a generative model ("hidden Markov model" in 18:35-42) which includes a set of model parameters ("alignment" in 18:35-42) that represent the entire image sequence ("entire video" in 18, line 37).

Regarding **claim 19**, Foote discloses the system of claim 1 further comprising a generative model which includes a set of model parameters that represent the images of the image sequence processed to that point (21:4-15).

Regarding **claim 22**, Foote discloses the system of claim 19 further comprising automatically reconstructing a representation of the image sequence from the generative model, wherein the representation comprises the preferred number of object classes (fig. 2, item 207).

Regarding **claim 23**, Foote discloses a computer-implemented process for automatically generating a representation of an object in at least one image sequence (fig. 1; fig. 2), comprising using a computer to:

acquire at least one image sequence (fig. 2, item 201), each image sequence having at least one image frame (fig. 3, items 301-308);

in near real-time automatically decompose each image sequence into a generative model (fig. 2, items 202-205; "Gaussian distributions" in 5, line 65-6, line 2), with each generative model including a set of model parameters (fig. 4; 7:59-60) that represent at least one object class for each image sequence using an expectation-maximization analysis (21:55-62) that employs a Viterbi analysis (6:43-45; 16:40-42).

Regarding **claim 24**, claim 2 recites identical features as in claim 24. Thus, references/arguments equivalent to those presented above for claim 2 are equally applicable to claim 24.

Claim Rejections - 35 USC § 103

[10] The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

[11] **Claims 4, 7,** and **27** are rejected under 35 U.S.C. 103(a) as being unpatentable over Foote et al. (US 6,404,925 B1) in view of Petrovic et. al (Transformed Hidden Markov Models: Estimating Mixture Models of Images and Inferring Spatial Transformations in Video Sequences, Computer Visions and Pattern Recognition, 2000, Vol. 2, pg 26 - 33).

Regarding **claim 4**, while Foote discloses the system of claim 3, Foote does not directly suggest wherein the model employs a latent image and a translation variable in learning each object class.

Petrovic discloses transformed hidden markov model wherein the model employs a latent image ("latent image", pg 27-28) and a translation variable ("set of transformations…", pg 27, right column) in learning each object class.

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the model of Foote to employ a latent image and a translation variable in learning each object class as taught by Petrovic to "develop a general video analysis tool that extracts long and short term similarities in video using a novel generative model, called the transformed hidden Markov model (THMM).", Petrovic, pg 26 and to "learn models of different types of object from unlabeled frames in a video sequence that include background clutter, occlusion and spatial transformations, such as translation, rotation and shearing.", Petrovic, pg. 26.

Regarding **claim 5**, while Foote discloses the system of claim 3, Foote does not directly suggest wherein the model describing the object class employs a latent image and a translation variable in filling in said hidden variables.

Petrovic discloses transformed hidden markov model wherein the model describing the object class employs a latent image ("latent image", pg 27-28) and a translation variable ("set of transformations…", pg 27, right column) in filling in hidden variables (pg 29).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the model of Foote to employ a latent image and a translation variable in filling in hidden variables as taught by Petrovic to "develop a general video analysis tool that extracts long and short term similarities in video using a novel generative model, called the transformed hidden Markov model (THMM).", Petrovic, pg 26 and to "learn models of different types of object from unlabeled frames in a video sequence that include background clutter, occlusion and spatial transformations, such as translation, rotation and shearing.", Petrovic, pg. 26.

Regarding **claim 27**, claim 4 recites identical features as in claim 27. Thus, references/arguments equivalent to those presented above for claim 4 are equally applicable to claim 27.

[12] **Claims 8-10, 13, 15-17,** and **28-31** are rejected under 35 U.S.C. 103(a) as being unpatentable over Foote et al. (US 6,404,925 B1) in view of Dellaert (<u>The Expectation Maximization Algorithm, College of Computing, Georgia Institute of Technology</u>, Technical Report number GIT-GVU-02-20, 2/2002).

Regarding **claim 8**, while Foote discloses a generalized expectation-maximization analysis, Foote does not directly teach wherein an expectation step of the generalized expectation-maximization analysis maximizes a lower bound on a log-likelihood of each image frame by inferring approximations of variational parameters.

Dellaert discloses the expectation maximization algorithm that teaches wherein an expectation step of the generalized expectation-maximization analysis maximizes a lower bound on a log-likelihood by inferring approximations of variational parameters (Section 2, "EM as Lower Bound Maximization").

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the generalized expectation-maximization for each image frame of Foote to include wherein an expectation step of the generalized expectation-maximization analysis maximizes a lower bound on a log-likelihood by inferring approximations of variational parameters as taught by Dellaert as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", Dellaert, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 9**, while Foote discloses a generalized expectation-maximization analysis, Foote does not directly teach wherein the maximization step of the generalized expectation-maximization analysis automatically adjusts model parameters in order to maximize a lower bound on a log-likelihood of each image frame.

Dellaert discloses the expectation maximization algorithm that teaches wherein the maximization step of the generalized expectation-maximization analysis automatically adjusts model parameters in order to maximize a lower bound on a log-likelihood (converting Θ into Θ^{t+1} in equation (4) in Section 2.2, "Maximizing the Bound").

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the generalized expectation-maximization for each image frame of Foote to include wherein the maximization step of the generalized expectation-maximization analysis automatically

adjusts model parameters in order to maximize a lower bound on a log-likelihood as taught by Dellaert as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", Dellaert, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 10**, while Foote discloses a generalized expectation-maximization analysis, Foote does not teach wherein the expectation step and the maximization step are performed once for each image in said image sequence.

Dellaert discloses the expectation maximization algorithm that teaches wherein the expectation step and the maximization step are performed once for each set of new data (equation (4) pg 6 to obtain Θ^{t+1} is only computed once for each set of new data).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for each image frame of the image sequence of Foote to be the new data as taught by Dellaert as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", Dellaert, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 13**, Foote discloses wherein automatic computation of the expectation step is accelerated by using a Viterbi analysis (6:43-45; 16:40-42; 18:31-48).

Regarding **claim 15**, while Foote discloses a generalized expectation-maximization analysis, Foote does not directly teach wherein the expectation-maximization analysis comprises: forming a probabilistic model having variational parameters representing posterior distributions; initializing said probabilistic model; inputting an image frame from the image sequence; computing a posterior given observed data in said image sequence; and using the posterior of the observed data to update the probabilistic model parameters.

Dellaert discloses the expectation maximization algorithm that teaches wherein the expectation-maximization analysis comprises:

forming a probabilistic model having variational parameters ("Θ^{t}", "Θ^{t+1}", means "θ_1" and "θ_2") representing posterior distributions (last paragraph, pg 1);

initializing said probabilistic model (the probabilistic model has to be initialized at some point to obtain Θ^{t+1});

inputting new data ("current guess" Θ^t from equation (3), pg 5 to "improved estimate" Θ^{t+1});

computing a posterior given observed data ("log-posterior log $P(\Theta|U)$", pg 6); and

using the posterior of the observed data to update the probabilistic model parameters ("M-step" equation, pg 6).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the new image frame from the image sequence of Foote to be the new data as taught by Dellaert and that the generalized expectation-maximization analysis of Foote to include wherein the expectation-maximization analysis comprises: forming a probabilistic model having variational parameters representing posterior distributions; initializing said probabilistic model; inputting; computing a posterior given observed data; and using the posterior of the observed data to update the probabilistic model parameters as taught by Dellaert as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", Dellaert, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 16**, Foote discloses wherein the expectation-maximization analysis further comprises:

outputting the model parameters (21:55-62).

Regarding **claim 17**, Foote discloses further comprising incrementing to the next image frame in said image sequence and repeating the actions after initializing the probability model until the end of the image sequence has been reached (the loops in fig. 12, fig. 20, fig. 26, and fig. 28 until frame sequence are complete).

Regarding **claim 28**, claim 8 recites identical features as in claim 28. Thus, references/arguments equivalent to those presented above for claim 8 are equally applicable to claim 28.

Regarding **claim 29**, claim 9 recites identical features as in claim 29. Thus, references/arguments equivalent to those presented above for claim 9 are equally applicable to claim 29.

Regarding **claim 30**, claim 15 recites identical features as in claim 30. Thus, references/arguments equivalent to those presented above for claim 15 are equally applicable to claim 30.

Regarding **claim 31**, claim 16 recites identical features as in claim 31. Thus, references/arguments equivalent to those presented above for claim 16 are equally applicable to claim 31.

[13] **Claims 11-12** are rejected under 35 U.S.C. 103(a) as being unpatentable over Foote et al. (US 6,404,925 B1) in view of Dellaert (The Expectation Maximization Algorithm, College of Computing, Georgia Institute of Technology, Technical Report number GIT-GVU-02-20, 2/2002) and Eberman et al. (US 5,925,065 A).

Regarding **claims 11** and **12**, while Foote in view of Dellaert disclose a computer-readable process of claim 8 wherein computation of the expectation step is suggested to use some form of transform, Foote in view of Dellaert does not teach accelerating the expectation step using a FFT-based inference analysis.

Eberman teaches using a FFT-based inference analysis (5:19-27).

It would have been obvious for the computation of the expectation step of Foote in view of Dellaert to include using a FFT-based inference analysis as taught by Eberman to reduce calculation time ($2N^2$) as less computation is needed ($2 N \log_2 N$) as well known to one of ordinary skill in the art.

It is well known to one of ordinary skill in the art that using the FFT requires performance on variables (x_n, k, N) that are converted into a coordinate system (X_k coordinate system) wherein transforms applied to those variables are represented by shift operations (x_n shifted by exponential on right side of equation to equal X_k).

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N}nk} \qquad k = 0, \ldots, N-1.$$

[14] **Claims 20-21** and **25-26** are rejected under 35 U.S.C. 103(a) as being unpatentable over Foote et al. (US 6,404,925 B1) in view of Jojic et al. (Learning Flexible Sprites in Video Layers, Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 2001, pg 1-8)

Regarding **claim 20**, while Foote discloses the system of 19, Foote does not teach wherein the model parameters include: a prior probability of at least one object class; and means and variances of object appearance maps.

Jojic discloses a learning flexible sprites in video layers wherein the model parameters include:

a prior probability of at least one object class ("prior probability *p(c)* of spring class *c*", pg 3); and

means and variances of object appearance maps ("means and variances of the sprite appearance maps", pg 3).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for system of Foote to include wherein the model parameters include: a prior probability of at least one object class; and means and variances of object appearance maps as taught by Jojic to "focus on learning the appearances of multiple objects in multiple layers, over the entire video sequence.", Jojic, pg 1 and to provide "probabilistic 2- dimensional appearance maps and masks of moving, occluding objects.", Jojic, pg 1.

Regarding **claim 21**, while Foote in view of Jojic discloses the system of 20, Foote in view of Jojic do not teach wherein the model further comprises observation noise variances.

Jojic discloses a learning flexible sprites in video layers wherein the model parameters include observation noise variances "the observation noise variances β", pg 3.

It would have been obvious to one of ordinary skill in the art at the time the invention was made for system of Foote to include wherein the model further comprises observation noise variances as taught by Jojic to "focus on learning the appearances of multiple objects in multiple layers, over the entire video sequence.", Jojic, pg 1 and to provide "probabilistic 2- dimensional appearance maps and masks of moving, occluding objects.", Jojic, pg 1.

Regarding **claims 25** and **26**, while Foote discloses the computer-implemented process of claim 23, Foote does not teach wherein the model parameters of each generative model includes

(i) an object class appearance map,

(ii) a prior probability of at least one object class, and

(iii) means and variances of that object class appearance map.

Jojic discloses a learning flexible sprites in video layers wherein the model parameters includes (i) an object class appearance map, (ii) a prior probability of at least one object class, and (iii) means and variances of that object class appearance map (Section 5, "Interference and Learning", first paragraph, pg 3).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for each generative model of Foote to include (i) an object class appearance map, (ii) a prior probability of at least one object class, and (iii) means and variances of that object class appearance map as taught by Jojic to "focus on learning the appearances of multiple objects in multiple layers, over the entire video sequence.", Jojic, pg 1 and to provide "probabilistic 2- dimensional appearance maps and masks of moving, occluding objects.", Jojic, pg 1.

[15] **Claim 32** is rejected under 35 U.S.C. 103(a) as being unpatentable over Foote et al. (US 6,404,925 B1) in view of Eberman et al. (US 5,925,065 A).

Regarding **claim 32**, claim 11 recites identical features as in claim 32. Thus, references/arguments equivalent to those presented above for claim 11 are equally applicable to claim 32.

Response to Arguments

[16] Applicant's arguments filed on February 25, 2008 with respect to claims 1-32 have been respectfully and fully considered, but they are not found persuasive.

[17] **Summary of Remarks** regarding **claim 1:**

Applicants argue that Foote does not teach the applicants' claimed preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time, nor does Foote teach in near real time automatically decomposing each image sequence into a generative model including a set of model parameters that represent at least one object class for each image sequence using an expectation-maximization analysis that employs a Viterbi analysis.

Applicant states that a "predefined set of classes" is not the same as a preferred <u>number</u> of classes, as the applicants claim. Cited Column 5, lines 14-16, does not teach "automatically decomposing the image sequence into the preferred number of classes of objects in <u>near real-time</u>." Nothing at all is stated in this paragraph regarding processing in near real-time. In fact, clearly Foote does not teach automatically decomposing the image sequence into <u>the preferred number</u> of classes of objects <u>in near real-time</u> because Foote segments <u>a full video into individual presentations based on the extent of each presenter's speech</u>. (Abstract) Hence, Foote can only segment a video file with corresponding audio after it has been recorded, not in real-time as it is being input. (Applicants' Resp. at 14-15, Feb. 25, 2008.)

[18] **Examiner's Response** regarding **claim 1**:

However, a "predefined-set of classes" is equivalent to a "preferred number of classes". A "set" is "a number of things of the same kind that belong or are used together." *See* <u>Merrian-</u>

Webster Online, 2007-2008, "set" n. def. 2, *available at* http://www.m-w.com/dictionary. A "pre-defined set" would be "preferred" as opposed to random, and if even random set of classes would have been preferred to be random. A "pre-defined set" is also a quantity, whether a whole number or the null-set. The examiner asserts that "predefined-set of classes" and "preferred number of classes" are highly equivalent. Foote el al. discloses a "pre-defined set of classes" at 5:14-16 for further use in fig. 12 (e.g., item 1204) which would require a "preferred number of classes of objects".

Similarly, "in near real-time" is again highly subjective as there does not any quantity or value assigned to the word "near" (e.g., "less than 10 seconds"). The adverb "near" is "at, within, or to a short distance in time". *See* Merrian-Webster Online, 2007-2008, "near" n. def. 2, *available at* http://www.m-w.com/dictionary. The Examiner contends two separate interpretations of the use of the phrase "in near real-time", both equally applicable.

First, the use of the phrase "in near real-time" is not specifically latched to any two separate events to constitute the time between them to be "in near real-time." "[A]utomatically decomposing the image sequence into the preferred number of classes of objects" by itself is "in near real-time" the instant it occurs in Foote et al. The instant

Second (even if the first argument fails), with respect to the entire age of the universe, the time between the two events argued by the Applicant of which would be "in near real-time", as even 10 days in comparison to the age of the universe would be "in near real-time."

The examiner suggests to further limit (to eliminate) such a broad interpretation of the claim in question (giving more definite language that would distinguish the application from the prior art of record), as both "preferred number of classes of objects" and "in near real-time" are highly

subjective. It is the Examiner's responsibility to interpret the claims as broad as possible and "preferred number of classes of objects" and "in near real-time" are undoubtedly anticipated by Foote et al.

[19] **Summary of Remarks** regarding **claim 23:**

Applicants argue that they have claimed an element not taught in Foote, namely inputting a number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time. Also Foote does not teach decomposing an image sequence into a generative model or decomposition of an image sequence inn near real time. (Resp. at 15.)

[20] **Examiner's Response** regarding **claim 23**:

However, the Examiner's response regarding claim 1 clarifies the broad interpretation of the elements in question.

[21] **Summary of Remarks** regarding **claims 4, 7,** and **27:**

Applicants argue, as discussed above Foote does not teach the applicants' claimed preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time, Nor does Foote teach in near real time automatically decomposing each image sequence into a generative model including a set of modem parameters that represent at least one object class for each image sequence using an expectation-maximization analysis that employs a Viterbi analysis. Petrovic also does not teach these features. (Resp. at 17-18.)

[22] **Examiner's Response** regarding **claims 4, 7,** and **27:**

However, the Examiner's response regarding claim 1 clarifies the broad interpretation of the elements in question.

[23] Summary of Remarks regarding **claims 4, 7,** and **27:**

Applicants argue, as discussed above Foote does not teach the applicants' claimed preferred <u>number</u> of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects <u>in near real-time</u>, nor does Foote teach <u>in near real time</u> automatically decomposing each image sequence into a generative model including a set of modem parameters that represent at least one object class for each image sequence using an expectation-maximization analysis that employs a Viterbi analysis. Dellaert also does not teach these features. (Resp. at 19.)

[24] Examiner's Response regarding **claims 4, 7,** and **27:**

However, the Examiner's response regarding claim 1 clarifies the broad interpretation of the elements in question.

[25] Summary of Remarks regarding **claims 11-12:**

Applicants argue, as discussed above Foote does not teach the applicants' claimed preferred <u>number</u> of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects <u>in near real-time</u>. Dellaert and Eberman also does not teach these features. (Resp. at 21.)

[26] Examiner's Response regarding **claims 11-12:**

However, the Examiner's response regarding claim 1 clarifies the broad interpretation of the elements in question.

[27] Summary of Remarks regarding **claims 20-21** and **25-26:**

Applicants argue, as discussed above Foote does not teach the applicants' claimed preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time, Nor does Foote teach in near real time automatically decomposing each image sequence into a generative model including a set of modem parameters that represent at least one object class for each image sequence using an expectation-maximization analysis that employs a Viterbi analysis. Jojic also does not teach these features. (Resp. at 22.)

[28] Examiner's Response regarding **claims 20-21** and **25-26:**

However, the Examiner's response regarding claim 1 clarifies the broad interpretation of the elements in question.

[29] Summary of Remarks regarding **claim 32:**

Applicants argue, as discussed above Foote does not teach the applicants' claimed preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time, Nor does Foote teach in near real time automatically decomposing each image sequence into a generative model including a set of modem parameters that represent at least one object class for each image sequence using an expectation-maximization analysis that employs a Viterbi analysis. Eberman also does not teach these features. (Resp. at 24.)

[30] Examiner's Response regarding **claim 32:**

However, the Examiner's response regarding claim 1 clarifies the broad interpretation of the elements in question.

Conclusion

[31] The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. US 5487117 A; US 5598507 A; US 5806030 A; US 6073096 A.

[32] **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

[33] Any inquiry concerning this communication or earlier communications from the examiner should be directed to DAVID P. RASHID whose telephone number is (571)270-1578. The examiner can normally be reached Monday - Friday 7:30 - 17:00 ET.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Vikkram Bali can be reached on (571) 272-7415. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system,

see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/David P. Rashid/
Examiner, Art Unit 2624

David P Rashid
Examiner
Art Unit 26244

/Vikkram Bali/

Supervisory Patent Examiner, Art Unit 2624